

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06019499 A**(43) Date of publication of application: **28.01.94**

(51) Int. Cl.

G10L 9/14(21) Application number: **04197416**(22) Date of filing: **02.07.92**(71) Applicant: **KOKUSAI ELECTRIC CO LTD**

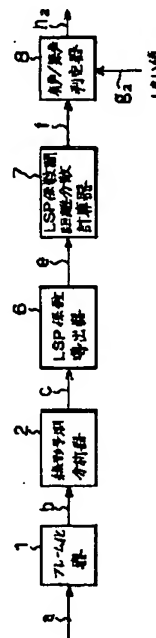
(72) Inventor:
SASAKI SEIJI
WATANABE OSAMU
GOTO HIROKI
MIYAKE MASAYASU

(54) VOICED/VOICELESS DECISION MAKING CIRCUIT**(57) Abstract:**

PURPOSE: To reduce the decision error rate of voiced/voiceless misdecisions on an input speech, made by an encoding device of an analytic synthesis system speech encoding system, due to a background noise.

CONSTITUTION: A linear predictive analyzer 2 finds a linear prediction coefficient (c) from a signal (b) framed by a framing unit 1 and an LSP coefficient deriving unit 6 converts the linear prediction coefficient (c) into an LSP(linear spectrum couple) coefficient (e); and then its variance is found by an LSP coefficient distance variance calculator 7 and a specific threshold value g_2 is supplied to a voiced/voiceless decision unit 8 to securely decide that the background is voiceless. Consequently, the reproduced speech quantity of a receiving and decoding device is improved.

COPYRIGHT: (C)1994,JPO&Japio-



'P

100-443887-100

1. The first part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

2. The second part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

3. The third part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

4. The fourth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

5. The fifth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

6. The sixth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

7. The seventh part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

8. The eighth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

9. The ninth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

10. The tenth part of the document is a list of names and addresses, which appears to be a directory or a list of contacts. The names are written in a cursive script, and the addresses are listed below them.

BLANK PAGE

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

2. Next, it is important to gather relevant information and data. This can be done through research, consultation with experts, or by analyzing existing data sets.

3. Once the information is gathered, the next step is to develop a plan or strategy to address the problem. This plan should outline the steps to be taken and the resources needed.

4. The fourth step is to implement the plan. This involves carrying out the tasks outlined in the plan and monitoring progress as it goes.

5. Finally, it is important to evaluate the results of the implementation. This involves comparing the actual outcomes with the expected outcomes and identifying any areas for improvement.

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平6-19499

(43)公開日 平成6年(1994)1月28日

(51)IntCl.⁵

G10L 9/14

識別記号

庁内整理番号

FI

技術表示箇所

C 8946-5H

審査請求 未請求 請求項の数2(全 6 頁)

(21)出願番号 特願平4-197416

(22)出願日 平成4年(1992)7月2日

(71)出願人 000001122

国際電気株式会社

東京都港区虎ノ門2丁目3番13号

(72)発明者 佐々木 誠司

東京都港区虎ノ門2丁目3番13号 国際電

気株式会社内

(72)発明者 渡辺 治

東京都港区虎ノ門2丁目3番13号 国際電

気株式会社内

(72)発明者 後藤 裕樹

東京都港区虎ノ門2丁目3番13号 国際電

気株式会社内

(74)代理人 弁理士 大塚 学

最終頁に続く

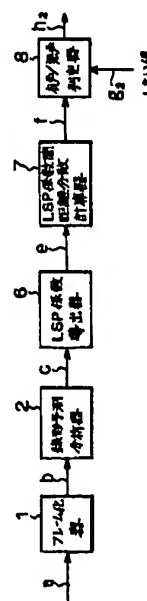
(54)【発明の名称】 有声/無声判定回路

(57)【要約】

【目的】分析合成系音声符号化方式の符号化装置における入力音声の有声/無声判定が背景雑音により誤判定となる判定誤り率を低減する。

【構成】フレーム化器1によってフレーム化された信号bを線形予測分析器2によって線形予測係数cを求め、その線形予測係数cをLSP係数導出器6によってLSP(線スペクトル対)係数eに変換した後、その分散をLSP係数間距離分散計算器7で求め、所定のしきい値g₂を有声/無声判定器8に与えて背景雑音を確実に無声と判定させるように構成した。

【効果】受信復号装置における再生音声品質が向上する。



【特許請求の範囲】

【請求項1】 フレーム化された音声信号を線形予測分析して線形予測係数を求める線形予測分析器と、該線形予測分析器からの線形予測係数を入力しスペクトル包絡の線スペクトル対係数を出力するLSP係数導出器と、

該LSP係数導出器から得られる線スペクトル対係数間の距離分散を出力するLSP係数間距離分散計算器と、該線スペクトル対係数間の距離分散を所定のしきい値と比較し、該分散がしきい値より大きいとき有声、小さいとき無声と判定してその結果を示す有声／無声フラグを出力する有声／無声判定器とを備えた有声／無声判定回路。

【請求項2】 入力音声信号をフレーム化するフレーム化器と、該フレーム化された音声信号を線形予測分析して線形予測係数を出力する線形予測分析器と、該線形予測係数をフィルタの係数として用い前記フレーム化された音声信号から予測残差信号を求める線形予測フィルタと、該予測残差信号を入力して時間遅れなしのときの相関値 R_0 と前記入力音声信号のピッチ周期 τ_p に相当する時間遅れのときの相関値 $R_{\tau p}$ とを出力する相関値計算器と、該2つの相関値の比 $R_{\tau p}/R_0$ を所定のしきい値と比較してその大小により有声／無声を判定して有声／無声フラグを出力する第1の有声／無声判定器とを備えた有声／無声判定回路において、前記線形予測分析器から得られる前記線形予測係数を入力しスペクトル包絡の線スペクトル対係数を出力するLSP係数導出器と、

該LSP係数導出器から得られる線スペクトル対係数間の距離分散を出力するLSP係数間距離分散計算器と、該線スペクトル対係数間の距離分散を所定のしきい値と比較し、該分散がしきい値より大きいとき有声、小さいとき無声と判定してその結果を示す有声／無声フラグを出力する第2の有声／無声判定器と、前記第1の有声／無声判定器と前記第2の有声／無声判定器とからそれぞれ出力される判定結果が、両方とも有声フラグのとき有声フラグを出力し、両方またはいずれか一方が無声フラグのとき無声フラグを出力する判定器とを備えたことを特徴とする有声／無声判定回路。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は分析合成系音声符号化方式の符号化装置に用いられる有声／無声判定回路に関する。

【0002】

【従来の技術】 低ビットレート（2.4 kbps程度）の音声符号化装置には分析合成符号化方式が適用されるのが一般的である。最も代表的な分析合成符号化方式としてはLPCボコーダと呼ばれる線形予測分析による音声符号化方式がある。この方式では、予測残差信号をパ

ルス列あるいは雑音にモデル化して情報圧縮している。入力音声は有声区間である場合はパルス列が用いられ、入力音声が無声区間である場合は雑音が用いられる。ここで重要なのは、符号化しようとしている音声フレームが有声区間であるか無声区間であるかの正確な判定である。この判定誤りが多発すると再生音声は聴感上で劣化が生じるので、判定誤り率を2%以下に抑える必要がある。

【0003】 従来技術としてLPCボコーダの有声／無声判定法について説明する。LPCボコーダでは、有声／無声の判定は音源（予測残差信号）の周期性の度合で決まる。周期性の度合は、入力音声信号のピッチ周期 τ_p に当たる時間遅れ（ $\tau = \tau_p$ ）での予測残差信号相関 $R_{\tau p}$ を、時間遅れなし（ $\tau = 0$ ）での予測残差信号相関 R_0 で正規化した値の大小で調べる。図4に時間遅れ τ と R_{τ}/R_0 （ R_0 で正規化された時間遅れ τ での予測残差信号相関 R_{τ} ）の関係を示す。ここで、有声／無声判定は次の条件式によって行われる。

$R_{\tau p}/R_0 \geq 0.25$ ならば 有声区間

$R_{\tau p}/R_0 < 0.25$ ならば 無声区間

図3は従来の有声／無声判定器の構成例図である。例えば8 kHzで標本化された入力音声aはフレーム化器1により1フレーム20 msecに分割されbとなりその後はフレーム毎に処理される。bは線形予測分析器2により線形予測分析され、その結果として線形予測係数cが得られる。cは線形予測分析フィルタ3の係数として用いられ、bは線形予測フィルタリングされ、予測残差信号dが得られる。dは相関値計算器4に入力され、時間遅れなし（ $\tau = 0$ ）での予測残差信号相関（ R_0 ）および入力音声信号のピッチ周期 τ_p に当たる時間遅れ（ $\tau = \tau_p$ ）での予測残差信号相関（ $R_{\tau p}$ ）が計算されて出力される。この2つの相関値 R_0 、 $R_{\tau p}$ は有声／無声判定器15に入力され上記の条件に基づき有声／無声判定が実行される。ここで、経験から求められるしきい値 g_1 （例えば0.25）が与えられている。判定結果である有声／無声フラグ h_1 は、例えば有声ならば0、無声ならば1と設定される。

【0004】

【発明が解決しようとする課題】 上記従来の回路構成において、送話者が周囲の背景雑音の少ない場所、例えば静かな事務室などに居る場合は特に問題はないが、一般的には背景雑音の存在を無視することはできない。従来の方式で問題となるのは、有声／無声判定において音声中の音源である予測残差信号の周期性を調べているのみであり、音声の音色を表す情報であるスペクトル包絡については考慮していないことである。つまり、音声は特有のスペクトル包絡を有するため有声／無声判定にこれを用いられれば有効であるにもかかわらず、使用していないことである。このため、周期性を有する背景雑音が入力音声に混在している場合、背景雑音区間（音声とは異な

るスペクトル包絡を有する)が有声区間であると誤って判定されてしまう。実例として、航空機内で録音した音声(信号対雑音比15dB)に対して従来の方式により有声/無声判定した結果の有声/無声判定誤り率は21%と大きな値を示した。ここで、有声/無声判定誤り率とは、有声/無声判定が誤ったフレーム数(入力音声のサウンドスペクトログラムを自視することにより求める)を全入力音声フレーム数で除算したものである。この結果、再生音声には“ビリビリ”という異音が目立つようになる。この原因は、分析合成系のLPCボコーダでは人間の音声のピッチ周期範囲は2:5~20ms(50~400Hz)であるという事実に基づいてピッチ周期抽出を行っているので、たとえ背景雑音に周期性があったとしてもピッチ周期がこの範囲を逸脱するため分析合成が正常に行われないためである。この現象は、背景雑音区間だけでなく、音声信号中の無声区間についても同様に生じる。従って、より正確な有声/無声判定方法が求められている。本発明の目的は、有声区間と無声区間でそれぞれに適した音源を用いる分析合成系音声符号化方式における再生音声品質を、送話側の背景雑音が大きい場合でも向上させて、有声/無声判定誤り率を2%程度に近づけるようにした有声/無声判定回路を提供することにある。ここで、2%という値は、入力音声に背景雑音が混入していない場合の有声/無声判定誤り率についての一般的な目標値である。

【0005】

【課題を解決するための手段】請求項1記載の本発明の有声/無声判定回路は、フレーム化された音声信号を線形予測分析して線形予測係数を求める線形予測分析器と、該線形予測分析器からの線形予測係数を入力しスペクトル包絡の線スペクトル対係数を出力するLSP係数導出器と、該LSP係数導出器から得られる線スペクトル対係数間の距離分散を出力するLSP係数間距離分散計算器と、該線スペクトル対係数間の距離分散を所定のしきい値と比較し、該分散がしきい値より大きいとき有声、小さいとき無声と判定してその結果を示す有声/無声フラグを出力する有声/無声判定器とを備えたことを特徴とするものである。以下この請求項1記載の本発明の技術思想について説明する。本発明は、有声/無声判定において音声の音源の周期性を調べるのみではなく、音声の音色を表す情報であるスペクトル包絡について音声(有声区間および無声区間)と背景雑音の性質の違いを識別することにより、有声/無声判定を行う際、背景雑音区間を有声区間であると誤判定する判定誤りを低減するものである。本発明は、スペクトル包絡情報としてLSP(線スペクトル対)係数を考え、その係数間の距離の分散により有声/無声判定を行うようにしたことを目ざとする。ここで、LSP係数とは、線形予測係数から導出される係数で、線形予測係数と等価なスペクトル包絡情報を表す。この係数の特徴としては、実際の

音声のスペクトル包絡と直接的な関係があることである。

【0006】図5(A)に音声信号の有声区間でのスペクトル包絡とLSP係数の関係を示す。同図では、線スペクトルが立っている周波数($f_1 \sim f_{10}$)が示されているが、これら f_i がLSP係数である。この図からわかるように、スペクトル包絡のピーク(ホルマントと呼ぶ)の近傍でLSP係数間の距離 $d_i (=f_{i+1} - f_i)$ が小さくなり、このピークが鋭くなるほど d_i は小さくなる。これに対してスペクトル包絡の谷間の部分では距離 d_i は大きく現れる。本発明では、このようなLSP係数の性質を音声(有声区間および無声区間)と背景雑音の識別に利用する。図5(A)に示すように、音声信号の有声区間でのLSP係数の典型的な出方として3または4個のホルマントが生じ、音声の無声区間の場合も数個のホルマントが生じるため、LSP係数間距離 d_i のばらつき(分散)は大きくなる。これに対して、背景雑音の典型的なLSP係数の出方は図5(B)に示すようにスペクトル包絡が平坦(ホルマントが生じない)なため、LSP係数間距離 d_i のばらつき(分散)が小さい。従って、 d_i の分散を計算し、その分散の値をしきい値と比較して判別することにより、音声(有声区間および無声区間)と背景雑音の識別ができる。しきい値は、予め音声および背景雑音でのLSP係数間距離 d_i の分散を観測することにより求められる。この詳細については第1の実施例で説明する。上述の手段により、背景雑音区間を有声区間であると誤判定する判定誤りを低減することができる。

【0007】次に、請求項2記載の本発明の有声/無声判定回路は、入力音声信号をフレーム化するフレーム化器と、該フレーム化された音声信号を線形予測分析して線形予測係数を出力する線形予測分析器と、該線形予測係数をフィルタの係数として用い前記フレーム化された音声信号から予測残差信号を求める線形予測フィルタと、該予測残差信号を入力して時間遅れ τ_p のときの相関値 R_0 と前記入力音声信号のピッチ周期 τ_p に相当する時間遅れのときの相関値 R_{τ_p} とを出力する相関値計算器と、該2つの相関値の比 R_{τ_p}/R_0 を所定のしきい値と比較してその大小により有声/無声と判定して有声/無声フラグを出力する第1の有声/無声判定器とを備えた有声/無声判定回路において、前記線形予測分析器から得られる前記線形予測係数を入力しスペクトル包絡の線スペクトル対係数を出力するLSP係数導出器と、該LSP係数導出器から得られる線スペクトル対係数間の距離分散を出力するLSP係数間距離分散計算器と、該線スペクトル対係数間の距離分散を所定のしきい値と比較し、該分散がしきい値より大きいとき有声、小さいとき無声と判定してその結果を示す有声/無声フラグを出力する第2の有声/無声判定器と、前記第1の有声/無声判定器と前記第2の有声/無声判定器とからそ

れぞれ出力される判定結果が、両方とも有声フラグのとき有声フラグを出力し、両方またはいずれか一方が無声フラグのとき無声フラグを出力する判定器とを備えたことを特徴とするものである。この請求項2記載の発明は、前記請求項1記載の発明に加えて、さらに、音声中の無声区間を有声区間であると誤判定する判定誤りを低減するものであり、前記本発明と予測残差信号相関を使用した従来方式を併用したものである。この方式では、無声区間の信号は、スペクトル包絡にホルマントが存在するため d_i の分散が大きくなり有声区間と誤判定されても、無声区間における音源は周期性を有していないため最終的には無声区間であると正確に判定される。この詳細については第2の実施例で説明する。

【0008】

【実施例】図1は本発明の第1の実施例を示すブロック図である。この実施例は、LSP係数間距離の分散を用いた有声/無声判定器の構成を示す。例えば、8kHzで標準化された入力音声aはフレーム化器1により1フレーム20msecに分割されbとなりその後はフレーム毎に処理される。bは線形予測分析器2により線形予測分析され、その結果として線形予測係数cが得られる。線形予測係数cはLSP係数導出器6によりLSP係数eに変換される。LSP係数eはLSP係数間距離の分散計算器7に入力され、分散fが得られる。有声/無声判定器8では、分散fとしきい値 g_2 (例えば1300Hz²)とを比較し、しきい値以上であれば有声、しきい値より小さければ無声と判定する。判定結果である有声/無声フラグ h_2 は、例えば有声ならば0、無声ならば1と設定される。上記のLSP係数間距離の分散計算器7内の処理は、例えば次のように行われる。線形予測分析器2内での分析次数を10次とすれば、LSP係数導出器6で導出されるLSP係数は図5のように10次となる。ここで、LSP係数間距離 d_i ($=f_{i+1} - f_i$) は $d_1 \sim d_9$ の9個が考えられる。LSP係数間距離の分散計算器7による分散の算出は次式で行われる。

【0009】

【数1】

$$\text{分散} = \frac{1}{n} \sum_{i=1}^n (d_i - \bar{d})^2$$

ただし、 \bar{d} は $d_1 \sim d_9$ の平均値であり、分散次数が10次の場合は $n=9$ となる。また、 d_i の単位はHzなので、分散の単位はHz² となる。このようにして算出された分散は、有声/無声判定器8でしきい値と比較される。従来方式と同様の入力音声(信号対雑音比15dBの航空機内で録音した音声)に対してこの方式に

より有声/無声判定した結果、有声/無声判定誤り率は10%となり従来方式に対し半減した。

【0010】図2は本発明の第2の実施例を示すブロック図である。この実施例は、LSP係数間距離の分散および従来方式(予測残差信号相関)を併用した有声/無声判定器の構成を示す。ここで、点線で囲まれたブロック10は従来方式(図3)と全く同じ構成であり、ブロック11は第1の実施例(図1)と全く同じ構成である。これらの有声/無声フラグをそれぞれ h_1 , h_2 とすれば、判定器9の動作は次のようになる。

(1) 最終的な結果である有声/無声フラグkを有声と設定する。

(2) もし、 h_1 , h_2 のうちどちらか、または両方が無声であれば h_3 を無声とする。そうでなければ、kは有声のままとする。

(3) kは、例えば有声ならば0、無声ならば1と設定される。

従来方式と同様の入力音声(信号対雑音比15dBの航空機内で録音した音声)に対してこの方式により有声/無声判定した結果、有声/無声判定誤り率は2.5%という値が得られ、従来方式に比べて大幅に低減された。

【0011】

【発明の効果】本発明を実施することにより、背景雑音の大きい場所から送話しても、有声/無声判定誤り率を、入力音声に背景雑音が混入していない場合の有声/無声判定誤り率についての一般的な目標値である2%程度に近づけることができるため、受信復号側の再生音声の品質向上に極めて大きい効果がある。

【図面の簡単な説明】

【図1】本発明の第1の実施例を示すブロック図である。

【図2】本発明の第2の実施例を示すブロック図である。

【図3】遅れ時間と予測残差信号の相関値の関係図である。

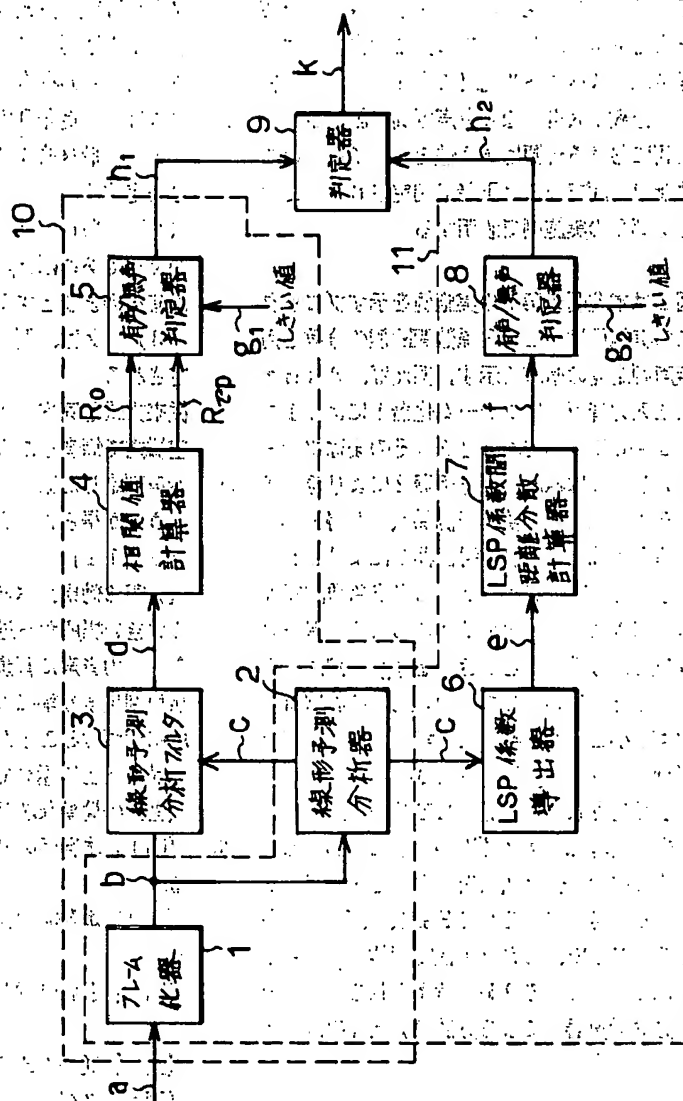
【図4】従来方式の構成例図である。

【図5】スペクトル包絡とLSP係数の関係図である。

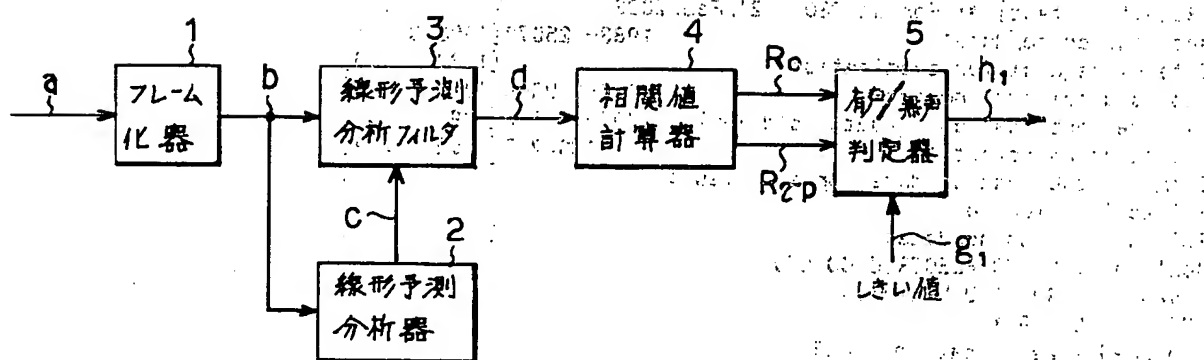
【符号の説明】

- 1 フレーム化器
- 2 線形予測分析器
- 3 線形予測分析フィルタ
- 4 相関値計算器
- 5 有声/無声判定器
- 6 LSP係数導出器
- 7 LSP係数間距離分散計算器
- 8 有声/無声判定器
- 9 判定器

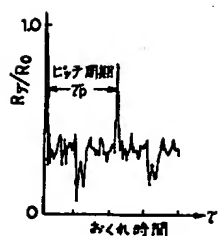
【図2】



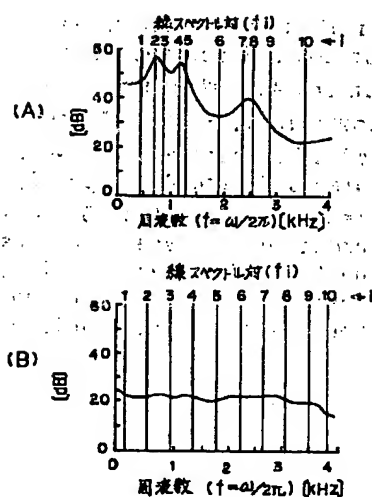
【図3】



【図4】



【図5】



フロントページの続き

(72)発明者 三宅 正泰

東京都港区虎ノ門二丁目3番13号 国際電
気株式会社内